DOCUMENT RESUME

ED 214 113 CS 006 529

AUTHOR Crismore, Avon
TITLE A Think-Aloud Protocol from a Critical Reader Reading a Study from "The Journal of Reading."
PUB DATE 82
NOTE 22p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Processes; *Critical Reading; *Periodicals; Reading Instruction; *Reading Research; *Reading Teachers; Secondary Education
IDENTIFIERS Protocol Analysis; *Think Aloud Protocol

ABSTRACT

Using the format of a think-aloud protocol, this paper critically reviews a study from "The Journal of Reading" from the perspective of a not-so-typical high school reading teacher. In this situation the hypothetical secondary school reading teacher came from an English major background, had 15 years experience teaching reading and English, subscribed to several professional journals, attended research sessions at regional conferences, and had had a few graduate courses in statistics. "The Journal of Reading" was chosen in order to gain insight into the nature of a research article in a nonresearch oriented journal. Another reason for the choice was its focus on secondary reading. The high school reading teacher perspective was chosen because reading educators and psychologists who write and submit research articles to the journal, as well as the editors and reviewers for the journal, need to know how a classroom teacher might react to these articles. Finally, the think-aloud technique was chosen because it related to the specific journal article being reviewed. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214113

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A Think-Aloud Protocol From a Critical Reader Reading a Study from The Journal of Reading

Avon Crismore

1982

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Avon Crismore

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS206529

The purpose of this paper is to critically review a study from The Journal of Reading from the perspective of a not-so-typical high school reading teacher, using the format of a think-aloud protocol. The Journal of Reading was chosen in order to gain insight into the nature of a research article in a non-research oriented journal. Another reason for this choice was its focus on secondary reading. The high school reading teacher perspective was chosen because reading educators and psychologists who write and submit research articles to the Journal of Reading and the editor and reviewers for the Journal of Reading need to know how a classroom teacher might react to these research articles. In this hypothetical situation, the secondary reading teacher comes from an English major background, has 15 years experience teaching reading and English, subscribes to The Journal of Reading and English Journal, occasionally reads The Reading Research Quarterly and Research in the Teaching of English, attends the research session at state and regional IRA and English conferences and has had one or two courses in statistics (courses not very well taught or learned) in graduate school. Such teachers do exist, but not in great quantity. The think-aloud technique was chosen because it relates to the specific study chosen, "The Effect of Two Contextual Conditions on Recall of a Reading Passage and on Thought Processes in Reading," and, just as importantly, because it might be more fun to write and also more interesting to read.

The not-so-typical reading teacher has just received the new Journal of Reading for April, 1981. The teacher decides to read the second

article, "The Effect of Two Contextual Conditions on Recall of a Reading Passage and on Thought Processes in Reading" since context and on-going thought processes in reading seem important issues and have been discussed in journal articles and conferences recently.

> The effect of two contextual conditions on recall of a reading passage and on thought processes in reading
>
> Christopherson teaches education at The University of Texas at San Antonio, Texas, Schultz at Trinity College, Hartford, Connecticut. Waern is with the Department of Psychology at the University of Stockholm, Sweden.
>
> Steven L. Christopherson
> Charles B. Schultz
> Yvonne Waern

Hmmm - looks like it will be about context, recalling, and processing. Sounds worthwhile. I wonder what the effect will be. I really didn't understand two contextual conditions. What are contextual conditions? What kinds of thought processes are there?

> Although various characteristics of text, such as context, organization, and difficulty, have been shown to affect recall in predictable ways, the thought processes involved in comprehension often remain unknown. Research methods rarely give us access to these processes, and in the absence of more direct evidence, researchers have frequently used the nature of observed recall, which occurs after reading, to make inferences about thought processes which were present during reading.
>
> A (new) research method using a think-aloud procedure, originated by Yvonne Waern (1979) at the University of Stockholm, illuminates many reading processes. The technique encourages readers to verbalize their thoughts as they read, usually at the end of each sentence. The readers' comments are recorded and subsequently are categorized according to the thought processes which are indicated.

They're going to use think-aloud protocol to illuminate the reading processes. Okay. Wait a minute! How can they call it a new research method. Jill Olshavsky used this method for looking at reading as problem

solving in an RRQ article in a 1976-1977 issue. And people in composing process research like Sondra Perl and Linda Flowers and John Hayes were using think-aloud techniques in 1979. Why weren't these studies cited? Don't these authors know the think-aloud methodology field or are they promoting Yvonne Waern? They said usually at the ends of sentences. Well, when don't the readers verbalize thoughts at the ends of sentences. Why did they decide to encourage the verbalization at the ends of sentences? What is the rationale? Why sentences? They didn't tell me. So - they're going to categorize the comments. I wonder how they'll do that. Will it be loose and sloppy or precise?

> The present study used the think-aloud procedure to elicit the thought processes associated with the contextual effects on recall found by Bransford and Johnson (1972).
>
> Bransford and Johnson had college students listen to a passage about washing clothes which was difficult to understand without its title, since the title contained the only clues to the concepts of clothing and washing. Seventeen students listened to the passage without knowing the title; 18 were given the title before listening to the passage. The students then wrote down what they could remember of the passage. The "title" group recalled significantly more than the "no title" group. Since both groups presumably had equal prior knowledge of washing clothes, the authors suggested that prior knowledge is helpful for comprehension and for recall only if it is activated, as in the condition with the title present.

There's that word contextual again. I don't know what it means. What are contextual effects? The authors described the Bransford and Johnson study simply and briefly. It seems easy to understand, but maybe I'd better read that study, too. How can Bransford and Johnson presume that their subjects all had equal prior knowledge of washing clothes. Well, by that time of their lives, maybe. In college today, students really do wash their own clothes. That's not true for high school students, though.

I've read this last paragraph and still no definition of contextual. Maybe I'll get it later.

Washing clothes

> The procedure is actually quite simple. First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run this may not seem important but complications can easily rise. A mistake can be expensive as well. At first the whole procedure will seem complicated. Soon, however, it will become just another facet of life. It is difficult to foresee any end to the necessity for this task in the immediate future, but then one never can tell. After the procedure is completed one arranges the materials into different groups again. Then they can be put into their appropriate places. Eventually they will be used once more and the whole cycle will then have to be repeated. However, that is part of life.
>
> [From Bransford and Johnson, 1972. Reprinted by permission of the authors.]

My students wouldn't understand another facet of life. Why would anyone want to read this passage? Would high school students want to read it? Would anyone really write a passage like that in the real world? It is unnatural--one of those experimented-constructed passages. Why didn't they use real texts?

> The think-aloud procedure which was used in the present study permits a more direct test of Bransford and Johnson's suggestion that readers with a context rely on prior knowledge more than readers without a context. The procedure also permits tests of hypotheses about thought processes during reading which were not considered by Bransford and Johnson.

The authors are assuming that think-aloud procedures are accurate--that self-reports are reliable. Well, maybe they will be more direct than the indirect recalls. Maybe.

> In our model of thought processes during reading, there are two sources of input for the reader. One is the current text and the other is prior knowledge, which includes world knowledge as well as information from prior text. The simplest processing which a reader may reveal to an observer is an attempt to represent the meaning of a segment of text such as a sentence. Once a representation is accomplished, the reader may attempt to match (or compare) the representation with prior knowledge. For example, when a reader comments, "That's true," there has been a successful match between the text information and the reader's own beliefs or prior knowledge. Sometimes the reader will attempt to construct new meaning after a representation has occurred. That is, the reader may combine information from the text with information from prior knowledge, as illustrated by the reader of a mystery novel who says, "Ah, I bet the butler did it."

Good--they give a model--the research seems grounded on a model. But I wonder what their model of reading is. Is a model of thought processes during reading the same as a model of reading? What is their model of comprehension? Top down? Bottom up? Interactive. What about world knowledge. World knowledge about what? It's too broad. About the topics? Does it have to be appropriate world knowledge to represent the meaning? Does world knowledge need to include knowledge about process writing (How to Wash Clothes), about style, tone, pragmatics? Does contextual include the sense of context of the situation? How do they define world knowledge? I don't like the author's examples. The comment "That's true" is a match between text and belief. Where is the example for a match of text and prior knowledge--world knowledge and prior text knowledge? Why did they use an example about a mystery novel to illustrate a new meaning construction when the authors used a nonfiction, process text? I want examples that relate to the text used in this study!

> Although our model assumes that a representation must precede a match or a construction, the representation often occurs so easily and automatically that the reader may only reveal an attempted match or construction. In contrast, a reader who has difficulty representing the meaning of a text is more likely to be aware of the attempts at representation and is less likely to be able to attempt matches and constructions. Representation is more difficult whenever relevant world knowledge is unavailable as, for example, in Bransford and Johnson's experimental condition where readers were not given a title and therefore had no meaningful context for the passage.

Okay--this all seems clear enough if you know the jargon. I wonder what a reader having difficulty representing the meaning of a text does to indicate to others and himself that he can't understand. Will these authors observe this behavior looking for patterns? Are there patterns for the same reader with different text types, different tasks, different situations? Well--now, context is related to no title. So that's what contextual effects are--title effects. Why didn't they explain this earlier? The authors didn't take me into consideration.

> The present study was designed to replicate Bransford and Johnson's study and to provide empirical evidence regarding which thought processes are associated with the two contextual conditions. Specifically, it was hypothesized that readers with a context (1) have greater recall than readers without a context, (2) make greater use of world knowledge, (3) are less likely to attempt consciously to represent the meaning of text sentences, (4) are more likely to attempt to match the text meaning with their prior knowledge, (5) are more likely to construct new meaning from the text and prior knowledge, and (6) are less likely to reread than readers without a context. The results verified all of these hypotheses except the last, where the anticipated difference was not statistically significant.

I understand the first purpose--it's clear. They'll replicate Bransford and Johnson's study. But why does it need to be replicated? Shouldn't they explain? Is there some problem with the original study?

I am curious about what readers do as they read either with or without a title (context). Why don't the authors just use the word title rather than context? Context confuses me. I don't understand their second hypothesis--make greater use of world knowledge. Seems too too fuzzy. Greater use--how? What kind of world knowledge? All kinds? Appropriate and related? Inappropriate and unrelated? I wonder--will the authors give examples and information about constructed new meanings? I hope so. That's important, because it is learning and that is what school is all about. Why did the author give me the results here? I thought that came later. Strange.

> Read/think plus recall
>
> The present study deviated from Bransford and Johnson's procedures in two important ways: Reading was substituted for listening and a think-aloud procedure was added. Thirty-five high school students volunteered to participate. Each student read the 15-sentence, 181-word passage on washing clothes which had been used by Bransford and Johnson.

Since this is a replication, and I am now interested in context effects and titles--I'll read the Bransford-Johnson article.

Some time later.

Well, I've skimmed through this study again since reading B-F, and this study now makes more sense. I'm thinking that if the think-aloud procedure is added the authors won't really be looking at reading. It's going to be a reading + situation. Being interrupted at the end of a sentence and discussing the thought processes is not a normal reading situation. The flow of comprehension is interrupted--short term and long term memory is now involved. How can they call it reading?

I'm angry! The authors tell me nothing about the high school subjects. Don't they know I'm interested in information about high school students because I teach them? What grade were they in high school? Were they good, average or poor readers? How smart are they? What kind of school do they attend--urban or rural, suburban--rich, poor? Are those students like my students? I've lost some confidence in the researchers. They don't understand my need to know this information--or is it the editor's fault? The students? Will this affect the study results? Will there be a Hawthorne effect?

> Twenty-five students read the passage without the title; 10 read it with a title. The more difficult task of reading the passage without a title was expected to increase the variability of performance; therefore, the majority of the students were given the more difficult task in order to reduce the error variance of the means of the dependent variables.

That seems dumb--25 read it without and 10 read it with the title. That's a strange proportion and it doesn't seem right. Why not a closer proportion like 20 without and 15 with the title? I'm not sure why they even want to reduce the error variance of the means of the dependent variables. I know having a lot of variability among the subjects is good.

> The experiment was conducted in individual sessions with each student. As a warm-up exercise for the think-aloud procedure, the students worked through a multiplication problem out loud. Then they read the experimental passage out loud and commented on their thoughts while reading.

They have 35 students reading one passage out loud. Why didn't they use fewer students and take a closer in depth look at their reading processes? The students could have been seen on 4 or 5 different occasions with different text types and tasks and context conditions?

How can they tell anything about high school readers with one experience with an artificial passage in an artificial reading situation? Why did they read it outloud? The authors assume oral and silent reading are alike. They aren't. The warm up exercise involved students working through a multiplication problem out loud. Multiplication problem! Why? Why not a reading passage similar to the experimental task?

I can about guess what those think-aloud protocols looked like, based on what I know about high schoolers. This whole procedure tries to externalize reading behavior, but it would be a difficult and distracting task for the subjects. I doubt that it really shows the actual inner processes. When they cannot, they are only approximately saying the first things that come to mind. Many things never get said, so the investigator must make inferences about what's going on underneath. The investigator maybe makes poor inferences! Perhaps with more sessions the think-aloud procedure would become more natural and comfortable, but with one session--No! How valid is one session? What about all that tacit knowledge the subject has that is never articulated? What about the lack of a schema for thinking aloud? What kind of comments did the subjects make? Any and all random thoughts that came to mind or selected ones? What direction did the subjects receive about commenting? The authors don't tell me enough.

> Finally, the students were asked to write down as much of the passage as they could recall. Most students required approximately 6 minutes for the whole task, with a range of about 4-15 minutes. Each student's recall was scored leniently for evidence of the original sentences.

Why was the recall scored leniently? What does that tell us about comprehension--why didn't they do two scorings? One lenient and one strict? Are the authors interested in comprehension? They used the word at the beginning of the article but have not since. If all they're looking for is evidence of the original sentences, any old thing would do. As long as the subject mentioned a noun or verb used in the original sentence it would count, no matter what the meaning of the sentence was and no matter whether it matched the meaning of the original. They should have done a propositional analysis of the passage and then of the recalls to see the matches and mismatches.

Now, wait a minute! This isn't a replication of the B-J study! In the original study, there were 3 conditions--No Topic. Topic After, and Topic Before. How come this study doesn't have 3 conditions? Why didn't the authors discuss this? The B-J study gave the standard error for the mean comprehension ratings and mean number of ideas recalled. Why didn't these authors do that? These authors have used passage B, a longer more informal passage than A. But passage B was used in Experiment III with 21 high school students, not college students. The college students were in Experiment II and used passage A and there were a total of 52--not 35 as the J of R study implies. The B-J study does not specify how many subjects were in the No Topic, Topic Before or Topic After condition. The Christopherson et al. group states there were 18 subjects given the title and 17 without a title. These authors are taking some information and procedures from Experiment II and some from Experiment III and combining it into a mythical experiment that B-J never did. The conditions

have been changed, the subjects changed, the passage changed, the task changed (no comprehension rating was used in the J of R study). The B-J study had different timing procedures than this study. In their study there was a minute interval between end of acquisition and comprehension rating, and 1 minute between the rating and recall. The recall was 6 minutes for all subjects in Experiment III with high school students. In Experiment II with college students, there was 2 minutes between end of acquisition and rating and 5 minutes allowed for recall. In the J of R study an average 6 minutes for recall was given with the range from 4 to 15 minutes. I don't know how long after oral reading and commenting the recall was done. With all the reading interruptions and commenting, many more minutes would pass between Sentence 1 and recall than in the B-J study and many distractions, so the recalls can't really be compared. How can this be a replication! It isn't!

> The reading and comments were tape recorded, and all coding was done from transcripts. Each reading and rereading of a sentence was coded as intake. The comments were first divided into simple sentences and independent clauses. Then each of these units was classified according to the following scheme, in which each comment (coding unit) is either related to the meaning of the text or is not related.

Nothing about the reliability of the coding. Was any inter-rater reliability done? The coding would be subjective. I wonder if the authors will be clear about their classification scheme. Will they have classification rules? Actual examples from student protocols?

> Some meaning-related comments indicate attempts at representation of the meaning of the text. An implicit or automatic attempt at representation, indicated by a comment such as "I see" or "I don't understand," was coded percept. A paraphrase or other explicit representation was coded interpretation. Other meaning-related comments may indicate attempts to match text meaning with prior knowledge, for

> example, "I didn't know that" or "That's not what the author said before." A reader may attempt to match current text with his or her world knowledge or with prior text. Comments may also indicate attempts at construction of new meaning from the text and prior knowledge. Construction may combine current text with world knowledge or with prior text, for example, "It's also the most dangerous mountain to climb" and "They must have dropped it from the bridge earlier." Three additional classifications are special cases where an attempt at construction leads to a statement of interest or disinterest, judgment of value, or free association.

I need more information about the categories and more examples from these subjects for this passage. The construction example is poor. Why use examples not related to the washing clothes passage? Why are there extra-text examples? Because there weren't any real constructions in this study? Why are the additional classifications a part of construction of meaning? These are not meaning-related comments in the cognitive sense. They are affective comments and need their own special category or else meaning needs to be defined broadly enough to include them. This is all vague and unclear.

> Some comments are non-meaning-related. Such comments may be about style. They may indicate thoughts which control processing, such as "I think I'll read that sentence again." Comments about the experimental task were coded meta.

Why isn't there more discussion about style? Example? I want more information! Why not more emphasis on control processing and meta comments? Why no examples for meta? Control processing--strategies are important! Attitudes and thought processes concerning style and meta are, also! This whole part frustrates me.

RESULTS

> Bransford and Johnson's finding that the presence of context improved the subjects' recall was replicated, in this case with reading rather than listening. The "context" and "no context" readers recalled a mean of 6.4 and 4.4 (out of 15) sentences respectively ($t(33) = 1.98$, $p < .03$, one-tailed).

A one tailed test? Why? Should they have? Is this supposed to impress me? It doesn't. I don't think it is really significant educationally. I don't agree. B-F's finding wasn't replicated. For Experiment II on recall, B-F had recall greater in Topic Before condition than for No Topic or Topic After conditions, $p > .005$ for both. In Experiment III, recall scores were higher in Topic Before than Topic After (there wasn't any No Topic condition) $p < .005$. That's quite a difference in significance between the present study and the B-F study. Perhaps this present study found differences due only to chance!

> The results of a multivariate analysis of variance are reported in Table 1. In addition to the codes described earlier, the dependent measures included the total number of remarks coded per protocol, excluding intake.
>
> The readers with no context did rely less on world knowledge than did readers with a context, as had been inferred by Bransford and Johnson; the means were 1.7 versus 6.1 items, respectively. The readers with no context also made many more conscious attempts to represent the meaning of the text than did readers with a context; the means were 11.6 versus 7.7. Readers with no context had significantly fewer matches and constructions than did the readers with a context, 1 versus 3.7 matches and 1.8 versus 3.7 construction.
>
> The one hypothesis which was not confirmed was that readers with no context would do more rereading. They did do twice as much rereading, 1.72 sentences as compared with .80 sentences for readers with a context, but the difference was not quite statistically significant.

They did use a multivariate analysis of variance, I see, which seems rather sophisticated. This doesn't jibe with the rest of the study. Much of it is unsophisticated. I'm puzzled. Is this a descriptive study?

Seems so from the classification of the think-aloud comments. But is is also statistical--here is analysis of variance--multivariate yet! A mixture! Why didn't the authors have discussion after each finding? They should have. Why didn't they tell me they were discussing the findings for their hypothesis? They assumed I'd know, huh? Why not use the words Hypothesis one? Why didn't they say what the numbers meant? Why isn't there a table for the recall data? It would be easier to interpret and compare with the B-J study.

Subjects' comments during reading:
Multivariate analysis of variance for the context and no context conditions

Direction of the comment	"Context" mean	"No Context" mean	F	P
Intake of text	15.80	16.72	2.75	.06*
Representation of text meaning	7.70	11.60	11.20	.001*
Percept	3.60	4.96	1.32	.26+
Interpretation	4.10	6.64	3.87	.06+
Match				
Current text and prior knowledge	3.70	1.00	11.60	.001*
Current text and world knowledge	3.50	.68	14.43	.001+
Current text and prior text	.20	.32	.23	.64+
Construction (includes interest, association, and judgment)	3.70	1.84	7.06	.006*
Current text and prior knowledge	1.90	1.16	2.01	.17+
Current text and world knowledge	.80	.32	2.01	.17+
Current text and prior text	1.10	.84	.40	.53+
Interest	.80	.16	7.45	.01+
Association	.70	.28	3.16	.09+
Judgment	.30	.24	.07	.80+
Nonmeaning-related comments	.70	1.32	1.57	.22+
Style	.60	.60	.00	1.00+
Control of processing	.00	.20	1.57	.22+
Meta (the experimental task)	.10	.52	2.75	.11+
Number of codes, excluding intake	15.90	15.92	.00	.99+

All classifications above which involve:				
World knowledge				
Including interest, association and judgment	6.10	1.68	19.65	.001*
Excluding interest, association and judgment	4.30	1.00	14.10	.001*
Prior text	1.30	1.16	.07	.81+

* one-tailed
\+ two-tailed

This table is ridiculous! I can't make heads or tails of it. Both one and two tailed texts! Crazy! Something hokey is going on. Why not one or the other?

I just don't like these dependent variables. The classification is wrong. It gives quantitative information based on frequencies, but I want qualitative information. Why didn't they control somehow the quality of the comments? It appears that the No Context subjects tried harder to make use of current text and prior text (.32) in matching than the context subjects with (.20). What does that imply about context? Why no discussion on this? Construction data seems misleading. If the interest, association and judgment is not included under construction, the no context subjects constructed (.48) meanings and the context subjects only (.10) meanings. What does this imply about context? Maybe students read more carefully without context or with more effort? It needs to be discussed. The no context subjects had (.52) for meta but the context subjects only (.10). I want to know the substance of their comments about meta. What is the effect on the findings? What kind of world knowledge appeared in the comments? I want to know. The table seems too sophisticated for the readers of J of R like me. Why wasn't it simplified?

Discussion

In addition to replicating Bransford and Johnson's finding of the effect of context on improving recall, this study found empirical evidence for the effect of context on the use of world knowledge, which had been hypothesized earlier but tested only indirectly. The think-aloud procedure which permitted the verification of the thought process hypothesized by Bransford and Johnson also permitted the verification of new hypotheses about the reading processes which are associated with the two different contextual condition.

The difference in recall between the context and no context conditions was statistically significant but not of the magnitude found by Bransford and Johnson. When listening, participants in the context condition had approximately a 100% advantage in recall over those in the no context condition, whereas when reading, those in the context condition had only a 50% advantage. Conscious rereading and unconscious visual regression may have enabled readers to reduce the disadvantage of being in the no context condition. More sophisticated readers, such as college students, may be able to reduce or eliminate the disadvantage in recall for the no context condition, but the striking differences in thought processing are not likely to disappear.

The discussion is separated from the results. This bothers me. I think the authors tried to manipulate figures so that their findings appeared statistically significant. I don't understand the sentence more sophisticated readers . . . for the no context condition. What is it supposed to mean? It's not clear.

> The findings to date suggest that some students who have problems understanding and recalling text may not have deficiencies in their listening and reading skills; their difficulties in comprehension and recall may be attributable to the unavailability of relevant poor knowledge, either because the information is not known or because its relevance is not preceived.

The findings do not suggest anything of the kind! They are going beyond their data. Nothing in the study pertained to comprehension. I don't have any information from the study that makes me think the students understood what they read even if there was some kind of evidence of the

original sentence in their recall. Nothing in the study pertained to relevance of prior knowledge either world knowledge or prior text knowledge. This paragraph is asserting things that are no doubt true for poor readers but nothing in the study supports these statements. The authors would have had more to say if they had done a multivariate analysis of variance for each of the 15 sentences--more significant things to say. I wonder, too, why they didn't have another condition--a group without the think-alouds. They would be more credible to me then. The sentence unit is a problem, too. Maybe (probably), readers process a word at a time, or a small group of words. Have the authors read much about eye movements and sentence processing? No citations about it.

> The think-aloud technique appears to have potential for revealing thought processes during reading. The technique is still being refined as we gain experience from this experiment and others in progress. We are exploring, for example, the possible benefits of separating comments which indicate representations into three classifications according to the reader's expression of success, failure, or uncertainty about each representation. As we improve upon the balance between the conceptual validity and reliability of the classification scheme, we expect that the think-aloud technique will be useful for studying reading under a variety of textual conditions and a variety of reader characteristics.

References

Barr, Anthony, James Goodnight, and Jane Helwig. Statistical Analysis System; version 76.6 Raleigh, N.C.: SAS Institute, 1976.

Bransford, John D., and Marcia K. Johnson. "Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall," Journal of Verbal Learning and Verbal Behavior, vol. 11 (1972). pp. 717-26.

Waern, Yvonne. "Thinking Aloud during Reading: A Descriptive Model and Its Application." Reports from the Department of Psychology, University of Stockholm, no. 546, April 1979.

So that's it. I guess I agree--think-alouds do have potential for revealing thought processes about reading, but the uses of this technique have lots to learn yet.

Disappointment. The study didn't achieve its goals very well. The study had important purposes and questions but a poor design, poor relaying of information with some needed information omitted and some poorly written. The study was too broad--it should have zeroed in on one aspect of the comments. The implications for secondary reading teachers and instruction were omitted, a major fault of the study. What does all this context-no context stuff mean for me as a reading or content area teacher? What about the think-aloud technique--could I use this technique myself? Should I? There wasn't enough information given so that I could, if I did want to. If the Journal of Reading is going to publish research studies then it needs to consider the concerns of classroom teachers who read research articles. I'm not impressed with the quality of the study or the editorial decisions on what to include or not include in research articles published in Journal of Reading.

After Thoughts - Reflections of the not-so-typical teacher at a Hypothetical Interview after the Think-aloud.

Well, I do feel that what the authors were trying to do in the study was significant, providing readers with a context, whether its pictorial or a title or background information about the situation for the writing of the text or the reading of it is important. Trying to find out the thought processes during reading, both meaning related and non-meaning related is also important as well as looking at the effect of two different

contextual situations and mental thought processes during reading. The authors did not make a strong case before or after the results for potential contribution to the field. The think-aloud technique was not handled well and there were no extra refinements of the procedure.

What would I have done if I had done the study? Several things.

1. Had 4 groups adding a group without the think-aloud.
2. Analyzed each sentence with a multivariate analysis of variance.
3. Maybe changed the sentence unit to a phrase unit or word unit.
4. Had both lenient and strict scoring of recalls.
5. Done a propositional analysis of the passage and the recalls in order to match.
6. Focused more on comprehension.
7. Had a completely descriptive plus statistical at end or vice versa but not a mixed study.
8. Used silent reading.
9. Used natural texts.
10. Focused on relevant world knowledge--appropriate world knowledge.
11. Used examples from the student comments and provided sample protocols in an appendix.
12. Used fewer subjects and more texts, a variety of texts, more think-aloud practice--had different reading tasks, more sessions, coded observable behavior while reading; used longer texts.
13. Used both good and poor readers--separated out by a pre-study.
14. Used the subjects on an eye movement machine like George McConkie's.
15. Written a section on educational implication and suggestions for high school teachers.

16. Worked out a more precise classification system for comments.
17. Looked for patterns for the same student over different texts and tasks.
18. Worked more on the attitude of reader toward author, text, style, tone, register, task.
19. Used interrater reliability on coding and recall scoring.
20. Randomly assigned subjects to conditions.
21. Used a regular classroom teacher to do the study to prevent the Hawthorne effect.
22. Worked for quality of learning or comments, not just frequency.
23. Let a classroom teacher read the study after it was written for comments on readability, tables, educational implications and suggestions.

This is all highly idealized, of course. I'm sure I'd need to make compromises and weigh the costs of each decision carefully in the real world of kids and classrooms. Given the opportunity to carry out such a study, no doubt I would look at the present study less critically, more sympathetically. The hyper critic might become the hypo critic.